Web Appendix

Democratic Values and Institutions

Timothy Besley and Torsten Persson

A Micro-Foundations for Political Institutions

We begin by discussing two examples that outline possible microfoundations for interpreting our framework in Section 3 of the text as a model of democracy. Each example focuses on one of the two main aspects of democratic institutions, namely open and free elections of the executive, on the one hand, and constraints on the executive (once in power), on the other. Both examples are highly stylized and can be considerably generalized.

Checks and balances The first example more nearly captures *constraints on the executive.* Here, we imagine that (a representative of) the incumbent group has proposal power over how to split some (resource) rents x_t across the two groups. This proposal will always allocate all the rents to the incumbent group. Under autocracy – i.e., with $D_t = 0$ – this proposal just goes through and we have

$$u^I(x_t, 0) = x_t \text{ and } u^O(x_t, 0) = 0.$$

Under democracy, $D_t = 1$, then instead with some exogenously given probability $2q < 1$, the opposition group can reject the proposal and impose an equal split of the rents with $x_t/2$ to each group. The expected rent allocation is thus

$$u^I(x_t, 1) = (1 - q)\, x_t \text{ and } u^O(x_t, 1) = qx_t.$$

Altogether, we have

$$\Gamma(x) = qx = \gamma(x).$$

Open elections The second example more nearly captures *open recruitment of the executive.* Under autocracy, $D_t = 0$, a representative of the incumbent group faces no challenge for power (but there may still be costly protests) and safely remains to the next period. But under democracy, $D_t = 1$, this representative runs against a representative of the opposition group in a stochastic electoral contest. The incumbent candidate wins this contest with probability $1 - x_t$. Thus $x_t \in [0, 1]$ is the (relative) unpopularity of the incumbent leader. We normalize the value of winning (which captures some unmodeled policy advantage) to 1. With $D_t = 0$, we have

$$u^I(x_t, 0) = 1 \text{ and } u^O(x_t, 0) = 0.$$

With $D_t = 1$,we instead have

$$u^I(x_t, 1) = 1 - x_t \text{ and } u^O(x_t, 1) = x_t.$$

Altogether, we have

$$\Gamma(x) = x = \gamma(x).$$

B Democratic Values

In this section, we discuss a possible microfoundation for the democratic values that appear in (2) of Section 3 in the text.

The expression in (2) assigns an additional positive payoff if $D_t = 1$ and a negative one if $D_t = 0$. It also assumes that democratic values are universal rather than particularistic. That is, concerned citizens care about society-wide gains and losses from democratic rights, and not only those which accrue to other concerned citizens. Assuming the latter would be an alternative way to formulate the model and would tend to strengthen the main results.

The formulation in (2) can be derived from a reference-dependent social preference, with one reference point for gains r_g and one for losses r_l

$$\text{(B.1)} \quad S(r_g, r_l, D, x) = \chi \min\left\{u^O(x, D) - u^O(x, r_l), 0\right\} + \max\left\{u^O(x, D) - u^O(x, r_g), 0\right\}.$$

We set $r_g = 0$ and $r_l = 1$ so gains are measured relative to the worst institution and losses relative to the best – i.e., concerned citizens evaluate social affairs against an institutional benchmark. The idea of reference-dependent preferences is well-established, following Kahneman and Tversky (1979) and a range of psychological studies. Specifically, our formulation follows Loomes and Sugden (1982), where an individual experiences either regret or rejoice depending on her reference point for an outcome.

Applications of reference-dependent preferences to concrete phenomena are discussed, e.g., in Kahneman et al (1991). (Koszegi and Rabin 2006 give a more recent theoretical treatment of reference-dependent preferences.) The payoffs in (B.1) can be thought as reflecting a feeling of (in)justice among citizens, based on *societal* gains/losses relative to the outcomes under the alternative institution, which embody their views about the right kind of society. Democratic values are thus distinct from standard preferences, analogous to the distinction between acquisition utility and transactions utility, which can also reflect a sense of justice (Thaler 1999).

C Socialization

In this section, we show three possible microfoundations for the evolutionary model stated in Section 3 of the main text.

Basic socialization model We first consider a model with successive generations, which overlap only in so far as parents endow their children with values, as modeled in Besley (2015). Children have two parents and – to keep the population balanced – all pairs have two children. We also assume that all marriage matching is random.[1]

Children are socialized into having democratic values. For simplicity, we model socialization as resulting from a form of osmosis rather than strategic behavior by parents.[2] Two parents of the same type simply pass along the values associated with their common type. However, children whose parents have different types get their type depending on the expected utilities of being concerned with democratic values rather than passive. Let $\Delta(\mu)$ be the expected utility difference between these types – their relative fitness – when the proportion concerned in the population is μ. Moreover, let $\eta \in (-\infty, \infty)$ be a couple-specific idiosyncratic negative shock to this utility difference. Then, a child with mixed parentage becomes concerned with democratic values, if and only if $\eta \leq \Delta$.

We assume that η has a symmetric single-peaked distribution with c.d.f. K and p.d.f. k. This implies that a mixed-marriage child holds democratic values with probability $K(\Delta(\mu))$ at utility difference $\Delta(\mu)$. By the law of large numbers, this is also the proportion among those with mixed parentage. By definition, $K(\cdot)$ is monotonically increasing, and by symmetry $K(0) = 1/2$.

The evolution of democratic values becomes:

$$\mu_{t+1} = \mu_t + 2\mu_t(1-\mu_t)\left[K(\Delta(\mu_{t+1})) - 1/2\right]. \tag{C.2}$$

This corresponds to (4) with $\varsigma^{P,C} = \mu\left[K(\Delta(\mu_{t+1})) - 1/2\right]$ and $\varsigma^{C,P} = -(1-\mu)\left[K(\Delta(\mu_{t+1})) - 1/2\right]$.

Strategic socialization We now show that the key equation (4) can be derived from a model, in which matching is assortative and socialization is purposeful. This follows the approach of Cavalli-Sforza and Feldman (1981) as adapted by Bisin and Verdier (2001). Socialization would then have two parts:

1. Direct Socialization: A parent may directly socialize a child into being a concerned citizen, depending on parental effort.

2. Oblique Socialization: If this is unsuccessful, the child may become socialized by society at large becoming a concerned citizen with probability μ_t.

We focus on the case where marriages are perfectly assortative and each pair of parents has two kids. Let $e \in \{0, 1\}$ be the effort put into socializing kids as concerned at cost C. Also, let the probability of successful socialization depend on $e + \varphi$ where φ is a stochastic socialization shock uniformly distributed on $\left[-\frac{1}{L}, \frac{1}{L}\right]$. Then, we have:

$$\text{Prob[concerned: } e] = \frac{1}{2} + Le.$$

[1] For the results to go through, we only require that there is at least some element of random matching. With full assortative matching, there would be no socialization as all offspring would have parents of the same type.

[2] This makes the model simpler and does not fundamentally affect the insights compared to the strategic socialization model of Bisin and Verdier (2001). .

Finally, as in our canonical model, let η be an idiosyncratic shock to parental preferences. They now choose socialization effort:

$$e^* = \arg\max \left\{ \left(\frac{1}{2} + Le \right) [\Delta(\mu) + \eta] - Ce \right\}.$$

This defines a threshold

$$\hat{\eta} = \nu - \Delta(\mu),$$

where $\nu = C/L$ such that $e^* = 1$ if and only $\eta \geq \hat{\eta}$.

For the children of concerned parents, the probability of a child being socialized as concerned is $K(\Delta(\mu_{t+1}) - \nu)$. For those who are not directly socialized, the probability of oblique socialization into being concerned is $(1 - K(\Delta(\mu_{t+1}) - \nu))\, \mu_t$.

Adding these expressions, the overall probability that the kids of concerned parents are concerned is:

$$K(\Delta(\mu_{t+1}) - \nu) + (1 - K(\Delta(\mu_{t+1}) - \nu))\, \mu_t. \tag{C.3}$$

If a child is born to passive parents, we assume she is never directly socialized into being concerned. However, she is socialized as passive with probability $(1 - K(\Delta(\mu_{t+1}) - \nu))$. The fraction of such children who are obliquely socialized as concerned is therefore:

$$K(\Delta(\mu_{t+1}) - \nu)\, \mu_t. \tag{C.4}$$

The overall fraction of concerned citizens in the next generation becomes

$$\begin{aligned} \mu_{t+1} &= \mu_t \left[K(\Delta(\mu_{t+1}) - \nu) + (1 - K(\Delta(\mu_{t+1}) - \nu))\, \mu_t \right] + (1 - \mu_t) \left[K(\Delta(\mu_{t+1}) - \nu)\, \mu_t \right] \\ &= (\mu_t)^2 + 2(1 - \mu_t)\, \mu_t K(\Delta(\mu_{t+1}) - \nu), \end{aligned}$$

which corresponds to (4) with $\varsigma^{P,C} = \mu_t \left[K(\Delta(\mu_{t+1}) - \nu) - \frac{1}{2} \right]$ and $\varsigma^{C,P} = -(1 - \mu_t) \left[K(\Delta(\mu_{t+1}) - \nu) - \frac{1}{2} \right]$. The only difference is that costly effort of being socialized as passive reduces the probability of concerned citizens in the population relative to our basic model, which has $\nu = 0$. This is the special case when $C = 0$ – i.e., when the effort by parents into socializing their child is costless.

In this setting, the candidate for an interior steady state is:

$$\Delta(\hat{\mu}) = \nu,$$

but when $\Delta_\mu(\mu) \geq 0$ this is unstable and the basic thrust of the basic-model analysis goes through unscathed.

A replicator dynamics Suppose that concerned and passive citizens are two behavioral types in the population and that members of each young generation adopts their types to the relative success of the "cultural parents" they encounter. This kind of imitation will give rise to a standard replicator dynamics:

$$\begin{aligned}\mu_{t+1}-\mu_t &= \mu_t \frac{[(\text{Utility Concerned:}\mu_t)-(\text{Average Utility:}\mu_t)]}{(1+\chi)\,\gamma(\bar{x})+\rho\underline{c}} \\ &= \mu_t(1-\mu_t)\frac{[(\text{Utility Concerned:}\mu_t)-(\text{Utility Passive:}\mu_t)]}{(1+\chi)\,\gamma(\bar{x})+\rho\underline{c}},\end{aligned}$$

where we have chosen to normalize by the maximum utility gain from democratic institutions so that the relevant expressions is bounded in the unit interval. Let $\pi(x,\mu)$ be the probability that $D=1$ given $\{x,\mu\}$. This expression boils down to

$$\begin{aligned}\mu_{t+1}-\mu_t &= \mu_t(1-\mu_t)\frac{\int_{\underline{x}}^{\bar{x}}\left[\pi(x,\mu_t)\,\gamma(x)-(1-\pi(x,\mu_t)\,L(x,\lambda(x,\mu)))\right]dH(x)}{(1+\chi)\,\gamma(\bar{x})+\rho\underline{c}} \\ &= \mu_t(1-\mu_t)\frac{\Delta(\mu_t)}{(1+\chi)\,\gamma(\bar{x})+\rho\underline{c}}.\end{aligned}$$

This is a special case of (**??**) if

$$\begin{aligned}\varsigma^{P,C} &= \frac{\mu_t \max\{\Delta,0\}}{(1+\chi)\,\gamma(\bar{x})+\rho\underline{c}} \\ &\text{and} \\ \varsigma^{C,P} &= \frac{(1-\mu_t)\max\{-\Delta,0\}}{(1+\chi)\,\gamma(\bar{x})+\rho\underline{c}}.\end{aligned}$$

Then the tipping point for the dynamics would be $\Delta(\hat{\mu})=0$, which would be similar to our analysis. Moreover, as long as $\Delta_\mu(\mu)\geq 0$, the dynamics would be qualitatively the same as in the canonical model.

D Steps 2 and 3 and Proposition 1

In this section, we analyze the optimal fighting decisions by the incumbent and the opposition, define the equilibrium functions $V(x_t,\mu_t)$, $U(x_t)$ and $\lambda(x,\mu)$ mentioned in the text, analyze their properties, and prove Proposition 1.

Protests and payoffs – step 3 All citizens observe the level of fighting f chosen at step 2 and protest if the benefit exceeds the cost. Given (3), passive citizens never protest as their private benefit is always lower than the cost. Therefore, the only issue is whether concerned citizens find it worthwhile to protest, given the realization of c_t. To determine this, define a threshold $\hat{c}(\mu,f,x)$ from the condition

$$\mu p(f)\left[u^O(x,1)-u^O(x,0)+s(x_t,1)-s(x_t,0)\right]=\hat{c},$$

i.e., the expected benefit from protesting equals the cost of protesting. Using (1) and (2) in the text, we can rewrite this condition as:

$$\hat{c}\left(\mu, f, x\right) = \mu p\left(f\right)\left[2 + \chi\right]\gamma\left(x\right).$$

Note that $\bar{c} > \hat{c}\left(\mu, f, x\right)$ for all $x \in [\underline{x}, \overline{x}]$ by (3). If $\underline{c} \leq \hat{c}\left(\mu, f, x\right)$, there is an equilibrium where all concerned citizens protest when $c_t = \underline{c}$ and the probability of a protest is therefore ρ. It is straightforward to see that a larger share of concerned citizens, μ, and/or a higher gain to democracy, x, increases the incidence of protests, while more incumbent fighting, f, reduces it.[3]

Now consider what happens when $D_t = 0$. The expected payoff to the incumbent leader with his preferred institution is $u^I\left(x_t, 0\right) + \widehat{\lambda}\left(\mu, f\right)\Gamma\left(x\right) - wf_t$, where $\widehat{\lambda}\left(\mu, f\right) = [1 - \rho\mu p(f)]$ is the probability of successfully enforcing $D_t = 0$ when devoting f units of labor to fighting.[4]

With democracy $D_t = 1$, we can write the leader's payoff as

(D.5) $$\widetilde{U}(x_t, f_t) = u^I\left(x_t, 1\right) - wf_t,$$

which takes into account the fact that no protest occurs in this case.

Choice of f – step 2 There is no incentive to fight when $D_t = 1$ and hence the payoff function under democracy is

(D.6) $$U(x_t) = \ \mathrm{Max}_f \widetilde{U}(x_t, f) = u^I\left(x_t, 1\right).$$

With autocracy, i.e. $D_t = 0$, fighting increases (via $p(f)$) the probability that an occurring protest is successfully defeated. The maximized expected payoff of an incumbent under autocracy ($D_t = 0$) is

(D.7) $$V\left(x_t, \mu_t\right) = u^I\left(x_t, 0\right) + \max_{f \geq 0}\left\{\widehat{\lambda}\left(\mu_t, f\right)\Gamma\left(x\right) - wf\right\}.$$

Let $f^*(x, \mu)$ denote the optimal choice of fighting by the incumbent at stage 2 and define the survival function $\lambda(x, \mu) = \widehat{\lambda}\left(\mu, f^*(x, \mu)\right)$.

[3] There is always an equilibrium where nobody protests. This has the possibility that protests can occur as "sunspot" phenomenon. Here, we assume that the concerned citizens can coordinate on the protest equilibrium when it exists.

[4] This objective function supposes that a passive incumbent-group citizen chooses the level of fighting. If we instead supposed that the decisions were made to maximize the average payoff in the incumbent group, then this would weaken their willingness to fight. Moreover, it would add an additional complementarity between democratic values and institutions, since a larger group of concerned citizens in the incumbent group would imply fewer resources devoted to fighting.

Properties of the equilibrium payoff and survival functions If none of the concerned citizens protest then $f^*(x,\mu)=0$. Given (3) there exists $\tilde{\mu}$ such that

$$\gamma(x)\,\tilde{\mu}(x)\,p(0)\,[2+\chi]=\underline{c}.$$

For $\mu\geq\tilde{\mu}(x)$ all concerned citizens protest when $c=\underline{c}$ and given the condition on $p(f)$ as f goes to zero. In this case, $f^*(x,\mu)$ solves

$$-\rho\mu p'(f^*(x,\mu))\Gamma(x)-w=0$$

The implicit-function theorem implies that

$$\frac{\partial f^*(x,\mu)}{\partial\mu}=\frac{-p'(f^*(x,\mu))}{p''(f^*(x,\mu))\,\mu}>0 \tag{D.8}$$

and

$$\frac{\partial f^*(x,\mu)}{\partial x}=\frac{-p'(f)\,\Gamma'(x)}{p''(f)\,\Gamma(x)}>0. \tag{D.9}$$

Now, we can substitute $f^*(x,\mu)$ into $\hat{\lambda}(x,\mu,\cdot)$ to define the incumbent's expected probability of successful enforcing $D_t=0$ when fighting optimally:

$$\lambda(x,\mu)=\left[1-\rho\mu p(f^*(x,\mu))\right].$$

It follows that

$$\lambda_x(x,\mu)=\begin{cases}-\rho\mu p'(f^*(x,\mu))\frac{\partial f^*(x,\mu)}{\partial x}>0 & \text{if }\mu\geq\tilde{\mu}(x)\\ 0 & \text{otherwise.}\end{cases}$$

Assume that

$$\begin{aligned}\lambda_\mu(x,\mu) &= -\rho\mu p'(f^*(x,\mu))\frac{\partial f^*(x,\mu)}{\partial\mu}-\rho p(f^*(x,\mu))\\ &= -\rho\left[\mu p'(f^*(x,\mu))\frac{\partial f^*(x,\mu)}{\partial\mu}+p(f^*(x,\mu))\right]\\ &= -\rho\left[\frac{-\left[p'(f^*(x,\mu))\right]^2}{p''(f^*(x,\mu))}+p(f^*(x,\mu))\right],\end{aligned}$$

which is negative if $\log(p(f))$ is convex. Thus

$$\lambda_\mu(x,\mu)=\begin{cases}-\rho\left[\frac{-[p'(f^*(x,\mu))]^2}{p''(f^*(x,\mu))}+p(f^*(x,\mu))\right]<0 & \text{if }\mu\geq\tilde{\mu}(x)\\ 0 & \text{otherwise.}\end{cases} \tag{D.10}$$

Moreover, we can write

$$V(x,\mu) - U(x) = \Gamma(x)\lambda(x,\mu) - wf^*(x,\mu). \tag{D.11}$$

We can use this expression to derive

$$\frac{\partial\left[V(x,\mu) - U(x)\right]}{\partial x} = \Gamma'(x)\lambda(x,\mu) + \lambda_x(x,\mu)\Gamma(x) > 0 \tag{D.12}$$

and

$$\frac{\partial\left[V(x,\mu) - U(x)\right]}{\partial \mu} = \begin{cases} \lambda_\mu(x,\mu)\Gamma(x) < 0 & \text{if } \mu \geq \tilde{\mu}(x) \\ 0 & \text{otherwise.} \end{cases}$$

Hence we have shown that, as stated in the main text of Section 3, for all $\mu \in [0,1]$ and $x \in [\underline{x}, \overline{x}]$

1. *A higher x increases $\lambda(x,\mu)$ and $V(x,\mu) - U(x)$.*

2. *A higher μ decreases $\lambda(x,\mu)$ and $V(x,\mu) - U(x)$.*

Proof of Proposition 1 Assumption 1 stated in the text requires that

$$\Gamma(\underline{x})\lambda\left(\underline{x},\underline{\mu}\right) - wf^*(\underline{x},\underline{\mu}) = 0,$$

which will hold only if

$$\gamma(\underline{x})\underline{\mu}p\left(f^*(\underline{x},\underline{\mu})\right)[2+\chi] \geq \underline{c}.$$

Hence there is both citizen protest by all concerned citizens when $c_t = \underline{c}$ and $f^*(\underline{x},\underline{\mu}) > 0$. Moreover, since $\gamma(x)$ is increasing then citizens protest for all $\mu \geq \underline{\mu}$ and $x \geq \underline{x}$. The decision rule used by the incumbent is

$$D_t = \begin{cases} 0 & \text{if } V(x,\mu) - U(x) \geq 0 \\ 1 & \text{otherwise.} \end{cases} \tag{D.13}$$

Let $\mu^L = \underline{\mu}$ as defined in Assumption 1. Then, for all $\mu \leq \mu^L$ and $x \in [\underline{x}, \overline{x}]$, we will have $D(\mu, x) = 0$. Since $V(x,1) - U(x) < 0$ for all $x \in [\underline{x}, \overline{x}]$, there exists

$$V\left(\bar{x},\mu^H\right) - U(\bar{x}) = 0.$$

Since $V(x,\mu) - U(x)$ is increasing, it follows that $\mu^H > \mu^L$. And for for all $\mu \geq \mu^H$, and $x \in [\underline{x}, \overline{x}]$, $D(x,\mu) = 1$. Given that $f^*(\underline{x},\underline{\mu}) > 0$, $V(x,\mu) - U(x)$ is a continuous function of μ and x for all $\mu \in \left[\mu^L, 1\right]$ and $x \in [\underline{x}, \overline{x}]$. Thus, for $\mu \in \left[\mu^L, \mu^H\right]$, the intermediate value theorem implies that there must be a value $\hat{x}(\mu) \in [\underline{x}, \overline{x}]$ such that

$$V(\hat{x}(\mu), \mu) - U(\hat{x}(\mu)) = 0. \tag{D.14}$$

E Dynamic Stability

This final section discusses the dynamic stability of the model.

The signs of $\Delta_\mu(\mu)$ and $d\hat{x}(\mu)/d\mu$ To rule out a stable interior steady state below it is sufficient that $\Delta_\mu(\mu) \geq 0$. This, in turn, is the case if $d\hat{x}(\mu)/d\mu > 0$. To see this, use (5) to compute:

$$\text{(E.15)} \qquad \Delta_\mu(\mu) = \begin{cases} \int_{\underline{x}}^{\overline{x}} \gamma(x)\, dH(x) & \mu \geq \mu^H \\ -\int_{\hat{x}(\mu)}^{\overline{x}} L_\lambda(x, \lambda(x,\mu))\, \lambda_\mu(x,\mu)\, dH(x) + \\ \left[\gamma(\hat{x}(\mu)) + L(\hat{x}(\mu), \lambda(\hat{x}(\mu),\mu))\right] h(\widehat{x}(\mu)) \frac{\partial \widehat{x}(\mu)}{\partial \mu} & \mu \in \left[\mu^L, \mu^H\right] \\ -\int_{\underline{x}}^{\overline{x}} L_\lambda(x, \lambda(x,\mu))\, \lambda_\mu(x,\mu)\, dH(x) & \mu \leq \mu^L. \end{cases}$$

Because $L_\lambda > 0$ and $\lambda_\mu < 0$, a sufficient condition for $\Delta_\mu(\mu) \geq 0$ for all $\mu \in [0,1]$, is $\partial\widehat{x}(\mu)/\partial\mu > 0$.

Using the definition of $\hat{x}(\mu)$, we can show that this condition is satisfied, because

$$\text{(E.16)} \qquad \frac{\partial \widehat{x}(\mu)}{\partial \mu} = -\frac{\partial V/\partial \mu}{\frac{\partial[V(x,\mu)-U(x)]}{\partial x}} = -\frac{\lambda_\mu}{\frac{\partial[V(x,\mu)-U(x)]}{\partial x}} > 0.$$

The sign follows from the results in section D, which say that the numerator is negative while the denominator is positive.

Stability We now provide the basic argument as to why only the corner solutions for μ can be stable steady states of the model.

We require that any steady state, $\hat{\mu}$, has to be stable following a small perturbation to $\hat{\mu} \pm \nu$. To prove that only the extremal steady states are stable, we start from

$$\text{(E.17)} \qquad \mu_{t+1} - \mu_t = (1-\mu_t)\,\varsigma^{P,C} - \mu_t \varsigma^{C,P}.$$

Note that if $\Delta > 0$ for all $\mu \in [0,1]$ then $\varsigma^{P,C} > 0$ and $\varsigma^{C,P} \leq 0$ and (E.17) is positive so μ converges to one globally. The opposite is true if $\Delta < 0$ for all $\mu \in [0,1]$. Now consider the case where there exists $\hat{\mu}(\sigma)$ such that $\Delta(\hat{\mu}) = 0$. Then since $\Delta(\mu)$ is globally increasing for $\mu \in [0,1]$, then at $\Delta(\hat{\mu}) = 0$, we must have $\mu_{t+1} - \mu_t \geq 0$ for all $1 \geq \mu \geq \hat{\mu}$, while $\mu_{t+1} - \mu_t < 0$ for all $0 \leq \mu < \hat{\mu}$. The interior steady state is therefore unstable. Moreover as $\Delta(\mu)$ is globally increasing, we must have $\Delta(1) \geq 0 \geq \Delta(0)$. Hence

$$\mu_{t+1} - 1 + \nu = (1-\nu)\,\varsigma^{P,C} - \nu\varsigma^{C,P} > 0$$

$$\mu_{t+1} - \nu = \nu\varsigma^{P,C} - (1-\nu)\,\varsigma^{C,P} < 0$$

for small enough $\nu > 0$. This implies that the steady states at $\mu = 0$ and $\mu = 1$ are stable as required. ■